Speaker identification based on speech rhythm: the case of bilinguals

Dellwo, Volker ; Schmid, Stephan ; Leemann, Adrian ; Kolly, Marie-José ; Müller, Mathias

Abstract: Voices are highly individual. The present study investigated how temporal characteristics of speech can contribute to speaker individuality for the same speaker speaking in different languages. By now there is a large body of evidence showing that measures based on temporal characteristics of consonantal and vocalic interval durations show drastic within language variability that is to a high degree a result of between speaker variability ([1], [2], [4]). Here we present results from an experiment on L2 and bilingual Italian/German speakers. Our assumption was that if speaker idiosyncratic rhythmic characteristics exist, then they should be present across utterances from different languages produced by the same speaker.

Preferences:

Format: oral or poster (both ok)
Wish to be considered for publication in an edited volume: no

Speaker identification based on speech rhythm: the case of bilinguals

Volker Dellwo, Stephan Schmid, Adrian Leemann, Marie-José Kolly and Matthias Müller

Phonetics Laboratory, University of Zurich, Switzerland
volker.dellwo@uzh.ch

Keywords: Speech rhythm, speaker idiosyncratic features

1. INTRODUCTION

Voices are highly individual. The present study investigated how temporal characteristics of speech can contribute to speaker individuality for the same speaker speaking in different languages. By now there is a large body of evidence showing that measures based on temporal characteristics of consonantal and vocalic interval durations show drastic within language variability that is to a high degree a result of between speaker variability ([1], [2], [4]). Here we present results from an experiment on L2 and bilingual Italian/German speakers. Our assumption was that if speaker idiosyncratic rhythmic characteristics exist, then they should be present across utterances from different languages produced by the same speaker.

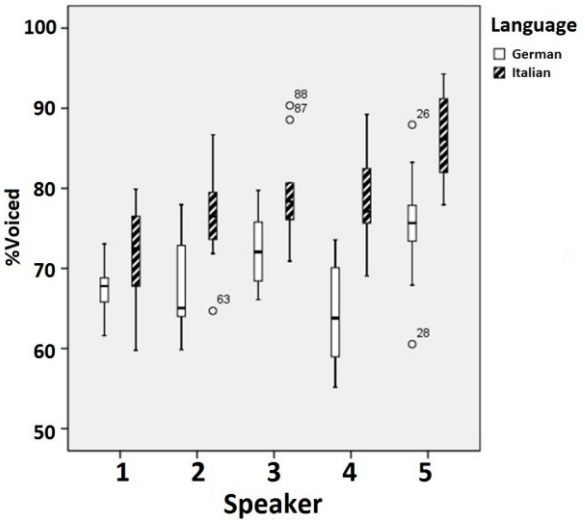
2. METHOD

5 Italian-German bilinguals, 5 German natives with L2 Italian and 5 Italian natives with L2 German were recorded reading 10 sentences (about 15 syllables on average) in each Italian and German. Speakers were selected to be similarly competent in their L2 across the two native language conditions (as judged by the second author). Durational characteristics of consonantal and vocalic intervals (e.g. syllable rate, %V, VarcoC and V, PVI, etc.) as well as voiced and unvoiced intervals (e.g. %Voiced, VarcoVoiced) were calculated for each sentence.

3. RESULTS & DISCUSSION

Results in figure 1 show the distributions (box-plots) of %Voiced (percentage over which speech is voiced) for each bilingual, speaking either Italian (striped plots) or German (white plots). Apart from the German condition of speaker 5 it can be observed that %Voiced has the tendency to increase from the leftmost to the rightmost speaker in both languages.

Figure 1: Box plot showing %Voiced for 5 bilinguals speaking German (white plots) and Italian (striped plots).



A factorial ANOVA (language * speaker) revealed that there is no interaction between the factors ($p=0.065$; the relatively low p-value is assumed to be caused by the German of speaker 4). However, highly significant main effects were found for both language ($p<.001$) and speaker ($p<.001$). The effects could be replicated for other metrics (e.g. VarcoV or syllable rate) and for the L2 speakers. Our results support the view that rhythmic characteristics of speech can be speaker-specific, independent of the language used by the speaker.

4. REFERENCES

[1] Dellwo, V. and Koreman, J. (2008) *How speaker idiosyncratic is measurable speech rhythm?* Abstract presented at IAFPA 2008, Lausanne/Switzerland (http://www.hf.ntnu.no/isk/koreman/Publications/2008/IAFPA2008abstract_DellwoKoreman.pdf)

[2] Loukina, A., Kochanski, G., Rosner, B. and Keane, E. (2011) *Rhythm measures and dimensions of durational variation in speech.* In: J. Acoust. Soc. Am. (129,5), 3258–3270.

[3] Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., and Mattys, S. (2010) *How stable are acoustic metrics of contrastive speech rhythm?* In: Journal of the Acoustical Society of America (127,3), 1559–1569.

[4] Yoon, T.J. (2010) *Capturing inter-speaker invariance using statistical measures of speech rhythm.* In: Electronic proceedings of Speech Prosody, Chicago/IL, USA.